Visualizing and Predicting Heart Diseases with an Interactive Dash Board

TEAM ID : PNT2022TMID19891

TEAM LEADER : 727819TUIT103 NAVANEETHAN C

TEAM MEMBER 1 : 727819TUIT124 RENITHA A

TEAM MEMBER 2 : 727819TUIT135 SHOBHA N

TEAM MEMBER 3 : 727819TUIT123 RAGUL N

TEAM MEMBER 4 : 727819TUIT148 VIBESHNAN P

PROJECT REPORT

1. INTRODUCTION

1.1. Project Overview

Cardiovascular diseases are the most common cause of death worldwide over the last few decades in the developed as well as underdeveloped and developing countries. Early detection of cardiac diseases and continuous supervision of clinicians can reduce the mortality rate. However, it is not possible to monitor patients every day in all cases accurately and consultation of a patient for 24 hours by a doctor is not available since it requires more sapience, time and expertise. Heart disease is perceived as the deadliest disease in the human life across the world. In particular, in this type of disease the heart is not capable in pushing the required quantity of blood to the remaining organs of the human body in order to accomplish

the regular functionalities. Heart disease include symptoms like physical body weakness, improper breathing, swollen feet, etc. The techniques are essential to identify the complicated heart diseases which results in high risk in turn affect the human life. As of, Healthcare industries generate enormous amount of data, so called big data that accommodates hidden knowledge or pattern for decision making. The huge volume of data is used to make decision which is more accurate than intuition. Exploratory Data Analysis (EDA) detects mistakes, finds appropriate data, checks assumptions and determines the correlation among the explanatory variables. In the context, EDA is considered as analyzing data that excludes inferences and statistical modelling. Analytics is an essential technique for any profession as it forecast the future and hidden pattern. Data analytics is considered as a cost effective technology in the recent past and it plays an essential role in healthcare which includes new research findings, emergency situations and outbreaks of disease. The use of analytics in healthcare improves care by facilitating preventive care and EDA is a vital step while analyzing data. In this project, the risk factors that causes heart disease is considered and predicted using interactive Dashboard and the analysis is carried out using a publicly available data in Kaggle for heart disease. The dataset holds 209 records with 8 attributes such as age, chest pain type, blood pressure, blood glucose level, ECG in rest, heart rate and four types of chest pain. To predict the heart disease, K-means clustering algorithm is used along with data analytics and visualization tool. The project discusses the pre-processing methods, classifier performances and evaluation metrics. In the result section, the visualized data shows that the prediction is accurate.

1.2 Purpose

Predicting and diagnosing heart disease is the biggest challenge in the medical industry and it is based on factors like physical examination, symptoms and signs of the patient. Here we use Exploratory Data Analysis (EDA) a method to analyze data using advanced techniques to expose hidden structure, enhances the insight into a

given dataset, identifies the anomalies and builds parsimonious models to test the underlying assumptions. Exploratory Data Analysis (EDA) is classified into Graphical or non-graphical and Univariate or multivariate data. Consider one data column at a time while multivariate method considers more than two variables while analyzing. Factors which influence heart diseases are cholesterol level of the body, smoking habit, and obesity, family history of diseases, blood pressure and working environment. Machine learning algorithms play a vital and accurate role in predicting heart disease. The advancement of technologies allows machine language to pair with big data tools to handle unstructured and exponentially growing data. In the project , K means clustering method and Linear Regression is proposed in big data environment and the data visualization is made with the help of IBM COGNOS and creating an Interactive Dashboard.

2. LITERATURE SURVEY

2.1 Existing problem

The existing solution uses the analysis which is carried out using a publicly available data for heart disease. The dataset holds 209 records with 8 attributes such as age, chest pain type, blood pressure, blood glucose level, ECG in rest, heart rate and four types of chest pain. The dataset is analyzed with visualization tool tableau and K means clustering. The dataset to define the proposed algorithm is the Cleveland heart disease raw dataset with 76 features of 303 patients. During the pre-processing method, some samples are removed to eradicate error due to inconsistency of data. The prediction of heart disease is made with 209 samples with seven independent features like age, chest pain type, blood pressure, blood glucose level, ECG in rest, heart rate and four types of chest pain and the habitual of physical exercise. Age is considered as the main risk factor for heart diseases as coronary fatty streaks develops in the adolescence stage. Male are at higher risk of coronary diseases than females, hence the data set considered here is for only male. Angina is the discomfort

caused when the muscles of heart is not supplied with sufficient oxygen rich blood. High blood pressure is one of the major causes of heart disease as it damages arteries. Blood pressure combined with diabetes can increase the risk even more. Heart rate with high blood pressure increases the risk of heart diseases. Heart beat rate is directly proportional to the risk of coronary disease. The symptom of heart disease includes feeling gripping and tight usually on the chest but spread to shoulders up to the stomach.

2.2 References

[1] Dr.A.V.Senthil Kumar, “Heart Disease Prediction Using Data Mining preprocessing and Hierarchical Clustering”, International Journal of Advanced Trends in Computer Science and Engineering, Volume-4, No.6, pp.07-18, 2015.

[2] Uma.K, M.Hanumathappa, “Heart Disease Prediction Using Classification Techniques with Feature Selection Method”, Adarsh Journal of Information Technology, Volume-5, Issue-2, pp.22-29, 2016.

[3] Himanshu Sharma, M.A.Rizvi, “Prediction of Heart Disease using Machine Learning Algorithms:A Survey”,International Journal on Recent and Innovation Trends in Computing and Communication,Volume5,Issue-8,pp.99-104, 2017.

[4] S.Suguna, Sakthi Sakunthala.N ,S.Sanjana, S.S.Sanjhana, “A Survey on Prediction of Heart Disease using Big data Algorithms”, International Journal of Advanced Research in Computer Engineering & Technology,Volume-6,Issue-3,pp.371-378,2017.

LINKS :

https://www.researchgate.net/publication/342621004

https://www.researchgate.net/publication/342621004_Heart_Disease_Prediction_using_Exploratory_Data_Analysis

2.2 Problem Statement Definition

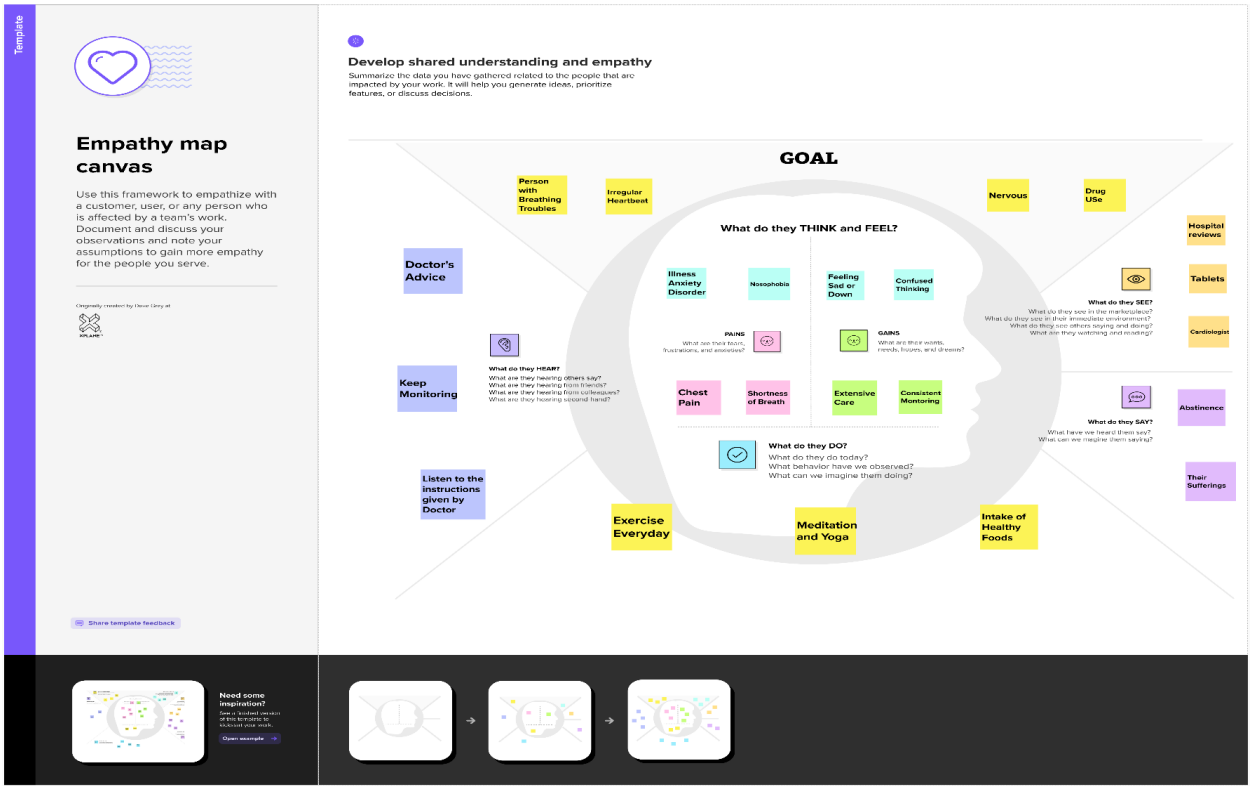
Blood arteries become blocked as a result of cardiac disease, which also causes the heart to stop beating. Numerous studies indicate that this condition has surpassed cancer as the leading cause of mortality. It's concerning because irregularities can only be found and acknowledged after the process has finished. However, if the illness is discovered quickly, it can be treated. The objective of this work is to create an interactive dashboard for visualizing and deconstructing cardiac disease. Based on the patient's medical characteristics, such as gender, age, chest discomfort, fasting sugar level, blood pressure, etc., this study seeks to determine if the patient is likely to be diagnosed with any cardiovascular heart disorders. A dataset containing the characteristics and medical background of the patient is chosen from the UCI repository. We have implemented this project to predict the heart disease of the patients having heart disease. The patients have to always consult the doctor whenever they get such symptoms and have to book an appointment whenever he consults the doctor. The patient have to wait in a queue to see the doctor. By this way the whole day goes waste. To overcome this issue we have bought this project into existence, where the time can be waste.The user has to enter their name, sex, blood pressure, maximum heart rate, blood sugar level and other details and by using those details we analyze and predict the risk of heart disease and the user gets benefited.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user’s behaviors and attitudes. Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user’s perspective along with his or her goals and challenges.

EMPATHY MAP



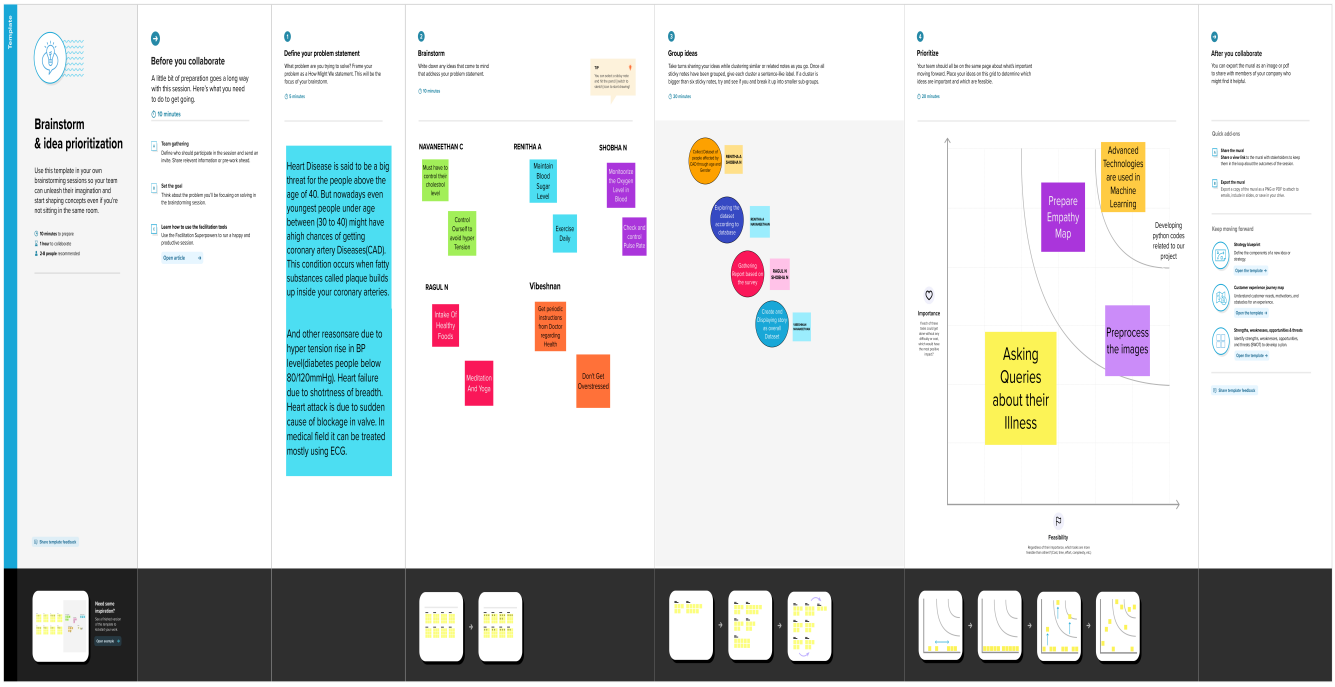
3.2 Ideation and Brainstorming

IDEATION :

Problem statement	I am (customer)	I'm trying to	But	Because	Which makes me feel
PS 1	a patient who has these symptoms in Chest such as chest-pain or some discomfort,Dizziness,Fainting (syncope) or near fainting,Fluttering in the chest Lightheartedness,Racing heartbeat (tachycardia), Shortness of breath,Slow heartbeat (bradycardia).	consult a doctor about the symptoms and want to know about the health status.In hospital they takes multiple tests to checkany disease that affects the heart .The test results also processed fora long amount of time .	after the multiple tests and after the long amount of time they give me the report that i have the disease called arrhythmia which is basically a heartbeat disease which occurs due to fast beat or slow beat of heart .I want to get rid of the disease but they said it become costly and they want more data to cure the disease.	the charges for taking the heart disease tests are very high and every one cannotafford it .There are large no of tests are taken and the time consumption alsobecome very high to get the results for the tests.Technology is developed in every field except medicalfield. we want to apply the technology in medical field alsoand want o lower the no of tests taken and want to lower the	In general, complications of heart arrhythmia may include stroke, sudden death and heart failure.The delay for publishing the results bring melots of pressure andover thinking and disturb my mental strength and i have the doubts like can i able to live int his earth.Some times i think about the money that i want to spend for taking thetests.It bring me a lot of fear that how can i able to manage economically.

				time taken to publish the results.	

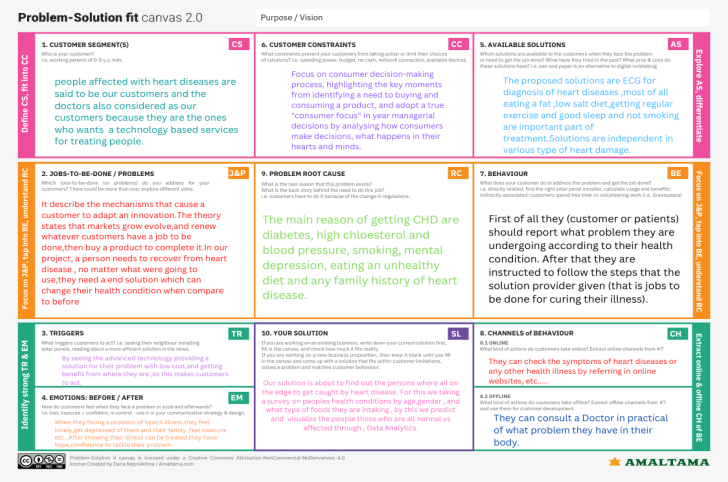
BRAINSTORMING:



3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to besolved)	1. To develop an interactive dashboard to predict the heart disease accurately with few predictability tests and attributes the presence of heart disease.
2.	Idea / Solution description	1. Analyzing data and identifying the heartdisease usingCognos analysis.
3.	Novelty / Uniqueness	1. Hoping to achieve maximum accuracy toprovide prior treatment to the patients and reduce the fatality rate.
4.	Social Impact/ Customer Satisfaction	1. Flexible and interactive dashboard. 2. Reduces the exorbitant medical cost of the patients. 3. Reduces the biases and mistakes caused by the decisions of doctors basedon their intuitions and experiences.
5.	Business Model(Revenue Model)	1. Ensure data security. 2. Ease of use. 3. Frequent updates based on the customer feedback.
6.	Scalability of the Solution	1. Can be used in any platform (Windows, mac, etc.,) 2. Updating with new feature is possible. 3. Scalable dataset.

3.4 Problem Solution Fit



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

FR No	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form. Registration through Gmail. Registration through Linked IN.

FR-2	User Confirmation	Confirmation via Email. Confirmation via OTP as message.
FR-3	User verification	Verification through CAPTCHA Verification through, I am not a robot.
FR-4	User Authentication	Recognition of correct person Resending the code in case of forgot password.
FR-5	User validation	Reconfirming the new password Sending a two digit number in (Google account) your Old devices, so that you can enter into a new device By entering the two digit number.
FR-6	User Submission	Submission through Google form Submission through Email.

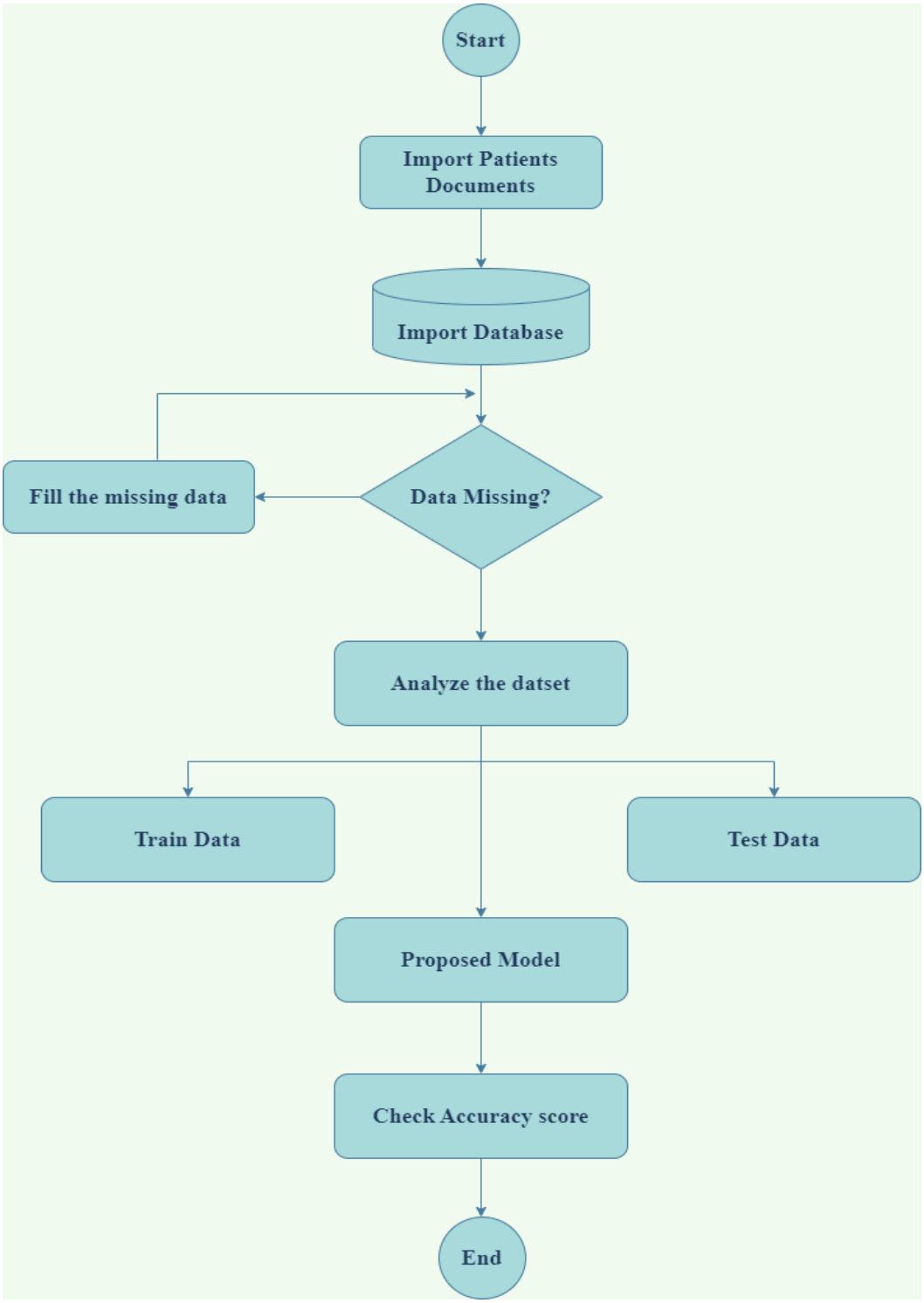
4.2 Non-Functional requirements

FR No	Non-Functional Requirement	Description
NFR-1	Usability	The EHDPS predicts the likelihood of patients getting heart disease. It enables significant knowledge, eg, relationships between medical factors related to heart disease and patterns, to be established.
NFR-2	Security	When it deals with(comes to)health factors, we should provide more security services. There shouldn't be no errors, lagging , base of data of a patient profile, while working on the software or product.
NFR-3	Reliability	Reliability is said to be the measure of stability or consistency of test scores shown in your product. Therefore your product will normal as a good performance one in the field of accuracy.
NFR-4	Performance	The performance should be fast relaying. This prediction system should be made available in cloud to ensure better accessibility and setting a milestone in providing good quality affordable healthcare.

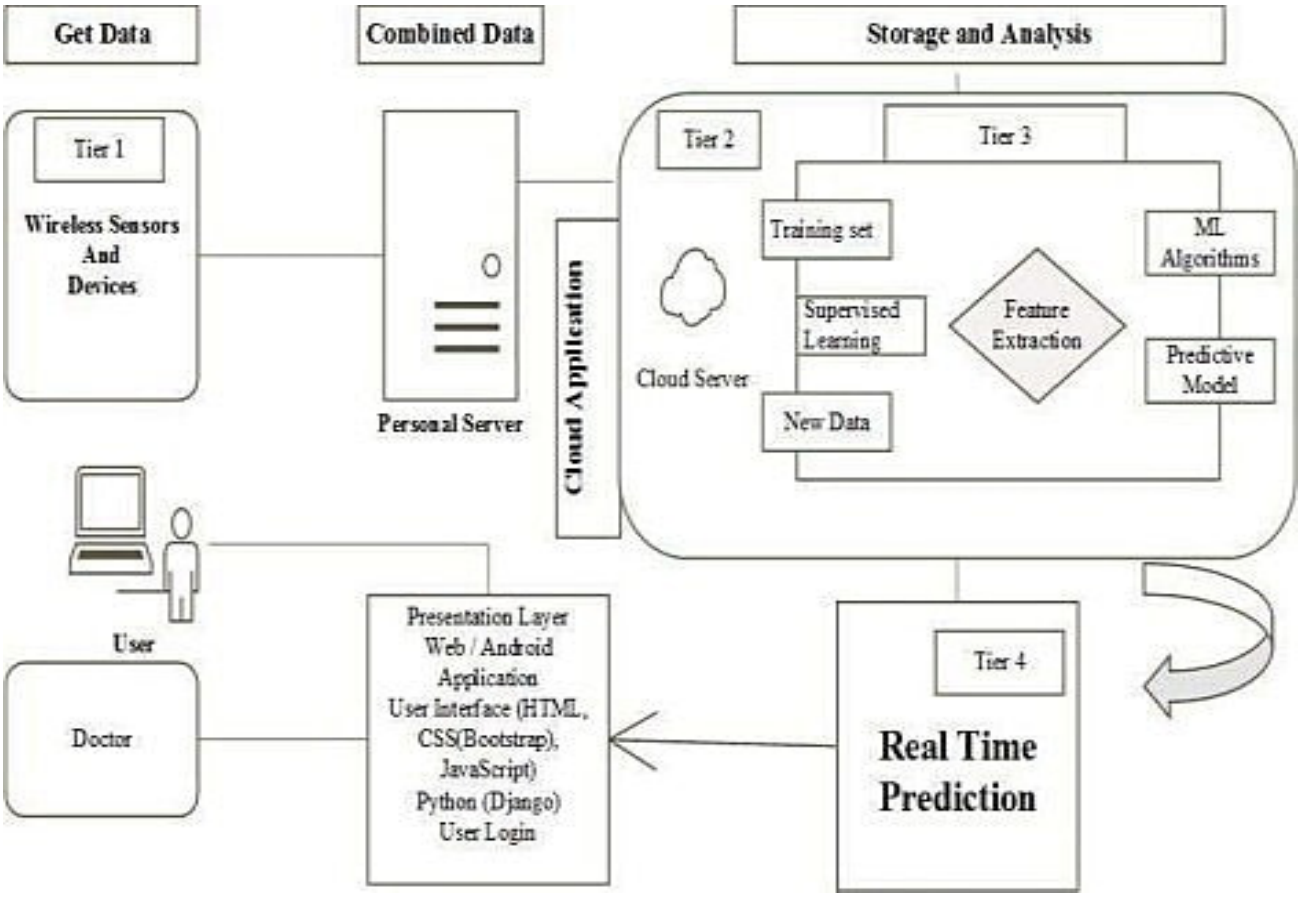
NFR-5	Availability	The Availability of getting used to this software or product design is through by accessing IBM cognos Analytics and IBM cloud.
NFR-6	Scalability	It is based on the number of users who maintaining the software or a system according to its performance like workflow, increase or decrease in efficiency , response time etc. Its scalability can be measured by maintenance, checking in for software updates, fixing errors if occurred in the server. By this a good quality of product is determined.

5. PROJECT DESIGN

5.1 Data Flow Diagrams



5.2 Solution & Technical Architecture



5.3 User Stories

User Type	Functional Requirement (Epic)	UserStory Number	User Story/ Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the	I can register & access the dashboard	Medium	Sprint-1

			application through Gmail	with Gmail Login		
	Login	USN-5	As a user, I can log into the application by entering email & password	I can register & access the dashboard with Gmail Login	High	Sprint-1
	Dashboard	USN-6	Profile - view & update your profile	I can see the profile.	Medium	Sprint-2
		USN-7	Change Password - user can change the password	I can able to change the password.	High	Sprint-1
		USN-8	Home - Analyze your Heart	I can detect the health condition from where ever I want	High	Sprint-1
		USN-9	The user will have to fill in the below 13 fields for the system to predict a disease -Age in Year -Gender -Chest Pain Type -Fasting Blood Sugar	These are the categories available in that application.	High	Sprint-2

			-Resting Electro graphic Results (Restecg) -Exercise Induced Angina(Exang) -The slope of the peak exercise ST segment -CA-Number of major vessels colored by fluoroscopy -Trest Blood Pressure -Serum Cholesterol -Maximum heart rate achieved(Thalac h) -ST depression induced by exercise(Oldpeak)			
		USN-10	View Doctors - view doctor detail by searching by names or filter by specialty	Using this application, people can known that the	Medium	Sprint-1

				sp ec ia l ty d oc to rs.		
Cus tom er (W eb use r)	System Requireme nt	USN-11	1. Hardware Requirement 1. Laptop or PC ■. I5 processor system or higher ■. 4 GB RAM or higher ■. 128GB ROM or higher 2. Android Phone (12.0 and above)	These are all the specificati on available on your PC.	High	Sprint-2
		USN-12	II. Software Requirement iii. Laptop or PC Windows 10 or higher Android Studio	Install your Applicati on. This system can be used to predict the presence of heart disease.	Medi um	Sprin t-2

		USN-13	Reference- https://ieeexplore .ieee.org/docume nt/9619208/	Go and Check our Reference link.	Medi um	Sprint-1
C us to m er Ca re	Dashboard	USN-14	Query	You can post your	High	Sprint-1

Executive				queries in the text box available in that application.		
		USN-15	Toll Free(8365492 10 7)	Ask your doubts in given number	High	Sprint-1
Administrator	Dashboard	USN-16	Verification	Verification through CAPTCHA Verification through I’m not a robot	High	Sprint-1

		USN-17	validation	Reconfirming the new password Sending a two digit number in (Google account) your Old devices, so that you can enter into a new device By entering the two digit number.	High	Sprint-2
		USN-18	Feedback - send feedback to the Admin.	Please send your feedback to the host.	Medium	Sprint-2

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

TITLE	DESCRIPTION	DATE
Literature Survey & Information Gathering	Literature survey on the selected project &gathering informationby referring to technical papers, research publicationsetc.	4 SEPTEMBER 2022
Prepare Empathy Map	Prepare Empathy Map Canvas to capture theuser Pains & Gains, Prepare list of problemstatements	8 SEPTEMBER 2022
Ideation	List them by organizing the brainstorming sessionand prioritize the top 3	18 SEPTEMBER 2022

	ideas based on feasibility & importance.	
Proposed Solution	Prepare the proposedsolution document, which includes the novelty, feasibility of idea, business model, social impact, scalabilityof solution, etc.	26 SEPTEMBER 2022
Problem Solution Fit	Prepare problem - solution fit document.	02 OCTOBER 2022

Solution Architecture	Prepare a solution architecture document.	1 OCTOBER 2022
Customer Journey	Prepare the customerjourney maps to understand the user interactions & experienceswith the application.	08 OCTOBER 2022
Data Flow Diagrams	Draw the data flow diagrams and submitfor review.	15 OCTOBER 2022
Technology Architecture	Architecture diagram.	16 OCTOBER 2022
Prepare Milestone & Activity List	Prepare the milestones &activity list of the project.	24 OCTOBER 2022
Project Development - Delivery of Sprint-1, 2, 3 & 4	Develop & submit thedeveloped code by testing it.	5 NOVEMBER 2022

6.2 Sprint Delivery Schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	1
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	2
Sprint-2		USN-3	As a user, I can register for the application through Facebook	2	Low	4
Sprint-1		USN-4	As a user, I can register for the application through Gmail	2	Medium	3

Sprint-1	Login	USN-5	As a user, I can log into the application by entering email &password	1	High	2
Sprint-2	Dashboard	USN-6	Profile - view & update your profile	2	High	5
Sprint-1		USN-7	Change Password - user can change the password	1	High	2
Sprint-1		USN-8	Home - Analyze your Heart	2	High	5

Sprint	**Functional Requirement (Epic)**	**User Story Number**	**User Story / Task**	**Story Points**	**Priority**	**Team Members**
Sprint-3		USN-9	The user will have to fill in the below 13 fields for the system to predict a disease -Age in Year -Gender -Chest PainType -Fasting BloodSugar -Resting Electro graphic Results(Restecg) -ExerciseInducedAngi	2	High	5

			na(Exang) -The slope of the peak exercise ST segment -CA – Number-of major vessels colored byfluoroscopy -Thal -Trest BloodPressure -Serum Cholesterol -Maximum heart rate achieved(Thalach) -ST depression induced by exercise(Old peak)			
		USN-10	View Doctors- view doctor detail bysearching by names or filter by specialty	1	Medium	4
Sprint-3	System Requirement	USN-11	A. Hardware Requirement i.Laptop or PC 1. I5 processor system or higher	2	High	2

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
			1. 4 GB RAM or higher 2. 128 GB ROM or higher ii. Android Phone (12.0 and above)			
Sprint-3		USN-12	II. Software Requirement iii. Laptop or PC 1. Windows 10 or higher 2. Android Studio	2	Medium	2
Sprint-4	Dashboard	USN-13	Query	1	High	1
		USN-14	Toll Free	1	High	1
		USN-15	Ratings	2	Medium	2
		USN-16	Verification	2	High	2
		USN-17	Validation	1	High	2
		USN-18	Feedback – send feedback to the Admin	2	Medium	3

Project Tracker, Velocity

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed(as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct2022	29 Oct2022	20	29 Oct2022
Sprint-2	20	6 Days	31 Oct2022	05 Nov2022	18	06 Nov2022
Sprint-3	20	6 Days	07 Nov2022	12 Nov2022	20	11 Nov 2022
Sprint-4	20	6 Days	14 Nov2022	19 Nov2022	19	19 Nov2022

7. CODING & SOLUTIONING

7.1 Feature 1

```
<!DOCTYPE html>
<html lang="en" dir="ltr">
  <head>
    <meta charset="utf-8">
    <title>Login</title>
    <link rel="stylesheet" href="styles.css">
<link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.2/dist/css/bootstrap.min.css"
rel="stylesheet" integrity="sha384-
Zenh87qX5JnK2Jl0vWa8Ck2rdkQ2Bzep5IDxbcnCeuOxjzrPF/et3URy9Bv1WTRi"
```

```
crossorigin="anonymous">
<script
src="https://cdn.jsdelivr.net/npm/bootstrap@5.2.2/dist/js/bootstrap.bundle.min.js"
integrity="sha384-
OERcA2EqjJCMA+/3y+gxIOqMEjwtxJY7qPCqsdltbNJuaOe923+mo//f6V8Qbsw3"
crossorigin="anonymous"></script>

 </head>
 <body >
  <nav class="navbar navbar-dark bg-dark">
    <div class="container-fluid">
 <a class="navbar-brand" href="#">Heart Disease Prediction</a>

  </div>
  </nav>
  <div class="row">
 <div class="col" style="background-color:white">
   <img src="heart.png" alt="" width="850" height="620" style="padding:3%">
 </div>
 <div class="col" style="background-color:white">
  <div class="login-form">
      <h1>Login Form</h1> <form action="home.html" method="post">
      <p>Email</p>
      <input type="text" name="mail" placeholder="Email">
      <p>Password</p>
      <input type="password" name="password" placeholder="Password"> <br>
      <button type="submit">Login</button>
      <p>Or <a href="signup.html">Sign Up</a></p>
    </form>
```

```
   </div>
  </div>
 </div>
 </body>
</html>
```

Style.css

```
 *{
            padding: 0;
            margin:0;
            font-family: sans-serif;
          }
          .login-form{
            width: 350px;
            top:50%;
            left:75%;
            transform:translate(-50%,-
          50%);
            position: absolute;
            color: black;
          }
          .login-form h1{
            font-size: 40px;
            text-align:center;
            text-transform:uppercase;
            margin: 40px 0;
          }
          .login-form p{
```

```
    font-size:20px;
    margin:15px 0;
  }
  .login-form input{
    font-size:16px;
    width:100%;
    padding:15px 10px;
    border:3px solid black;
    outline: none;
    border-radius:5px;
  }
  .login-form button{
    font-size:18px;
    font-weight: bold;
    margin:20px 0;
    padding:10px 15px;
    width:50%;
    border-radius:5px;
    border:0;
  }
  .row{
    width:100%;
  }
  .footer{
    width: 100%;
    height:115px;
    background-color: #2B2B2B;
}
```

7.2 Feature 2

```
<!DOCTYPE html>
<html lang="en" dir="ltr">
  <head>
    <meta charset="utf-8">
    <title></title>
    <link rel="stylesheet" href="styles.css">
    <link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.2/dist/css/bootstrap.min.css"
rel="stylesheet" integrity="sha384-
Zenh87qX5JnK2Jl0vWa8Ck2rdkQ2Bzep5IDxbcnCeuOxjzrPF/et3URy9Bv1WTRi"
crossorigin="anonymous">
    <script
src="https://cdn.jsdelivr.net/npm/bootstrap@5.2.2/dist/js/bootstrap.bundle.min.js"
integrity="sha384-
OERcA2EqjJCMA+/3y+gxIOqMEjwtxJY7qPCqsdltbNJuaOe923+mo//f6V8Qbsw3"
crossorigin="anonymous"></script>
  </head>
  <body >
    <nav class="navbar navbar-dark bg-dark">
      <div class="container-fluid">
  <a class="navbar-brand" href="#">Register</a>
</div>
    </nav>
    <div class="row">
   <div class="col" style="background-color:white">
     <img src="heart.png" alt="" width="850" height="620" style="padding:3%">
   </div>
   <div class="col" style="background-color:white">
```

```
    <div class="login-form">

        <form action="#" method="post">
        <p>Email ID</p>
        <input type="text" name="user" placeholder="Email">
        <p>Password</p>
        <input type="password" name="user" placeholder="Password">
        <p>Confirm Password</p>
        <input type="password" name="password" placeholder="Confirm Password"> <br>
        <button type="submit">Register</button>
        <p>Or Already have an account? <a href="index.html">Login</a></p>
     </form>
   </div>
  </div>
 </div>
 </body>
</html>
```

Style.css

```
*{
 padding: 0;
 margin:0;
 font-family: sans-serif;
}
.login-form{
 width: 350px;
 top:50%;
 left:75%;
 transform:translate(-50%,-50%);
```

```
  position: absolute;
  color: black;
}
.login-form h1{
  font-size: 40px;
  text-align:center;
  text-transform:uppercase;
  margin: 40px 0;
}
.login-form p{
  font-size:20px;
  margin:15px 0;
}
.login-form input{
  font-size:16px;
  width:100%;
  padding:15px 10px;
  border:3px solid black;
  outline: none;
  border-radius:5px;
}
.login-form button{
  font-size:18px;
  font-weight: bold;
  margin:20px 0;
  padding:10px 15px;
  width:50%;
  border-radius:5px;
  border:0;
```

```
}
.row{
  width:100%;
}
.footer{
  width: 100%;
  height:115px;
  background-color: #2B2B2B;
}
```

7.3 Feature 3

index.html

```
<!DOCTYPE html>
<html>
  <head>
    <title>Heart Disease Prediction</title>
    <link
href='https://fonts.googleapis.com/css?family=Open+Sans:400,300,300italic,400italic,6
00' rel='stylesheet' type='text/css'>
    <link href="https://fonts.googleapis.com/css?family=Roboto:300,400,500,700"
rel="stylesheet">
    <link href="style.css" rel="stylesheets">
  </head>
  <body>
    <img src="image.png">
    <div class="main-block">
    <form action="{{url_for('get_prediction')}}" >
      <h1>Hearth Disease Prediction</h1>
      <fieldset>
```

```
<legend>
  <h3>Personal Informations</h3>
</legend>
<div  class="account-details">
  <div><label>Name </label><input type="text" name="name" required></div>
  <div><label>Sex </label><select id="sex" name="sex">
    <option value="Male">Male</option>
    <option value="Female">Female</option>
    </select>
  </div>
  <div><label>Age </label>
    <select id="age" name="age">
    <option value="18-24">18-24</option>
    <option value="25-29">25-29</option>
    <option value="30-34">30-34</option>
    <option value="35-39">35-39</option>
    <option value="40-44">40-44</option>
    <option value="45-49">45-49</option>
    <option value="50-54">50-54</option>
    <option value="55-59">55-59</option>
    <option value="60-64">60-64</option>
    <option value="65-69">65-69</option>
    <option value="70-74">70-74</option>
     <option value="75-79">75-79</option>
    <option value="80 or older">80+</option>
    </select>
  </div>
  <div><label>Race </label>
    <select id="race" name="race">
```

```
          <option value="Asian">Asian</option>
          <option value="American Indian/Alaskan Native">American Indian/Alaskan
Native</option>
          <option value="Black">Black</option>
          <option value="Hispanic">Hispanic</option>
          <option value="White">White</option>
          <option value="Other">Other</option>
          </select>
        </div>
        <div><label>Height </label>
          <input type="number" id="height" name="height" placeholder="cm" min="0"
max="225" required>
        </div>
        <div><label>Weight </label>
          <input type="number" id="weight" name="weight" placeholder="kg" min="0"
max="225" required>
        </div>
      </div>
    </fieldset>
    <fieldset>
      <legend>
        <h3>General Information</h3>
      </legend>
      <div  class="personal-details">
        <div>
          <div><label>Have you smoked at least 100 cigarettes in your entire life?</label>
            <select id="smoking" name="smoking">
            <option value="Yes">Yes</option>
            <option value="No">No</option>
```

```
          </select>
        </div>
        <div><label>Heavy alcohol drinkers (Male: >14 drinks & Female: >7 drinks per
week)</label>
          <select id="alcohol" name="alcohol">
          <option value="Yes">Yes</option>
          <option value="No">No</option>
          </select>
        </div>
        <div><label>Would you say that in general your health is..</label>
          <select id="general_health" name="general_health">
          <option value="Excellent">Excellent</option>
          <option value="Very good">Very Good</option>
          <option value="Good">Good</option>
          <option value="Fair">Fair</option>
          <option value="Poor">Poor</option>
          </select>
        </div>
        <div><label>On average, how many hours of sleep do you get in a 24-hour period?
</label>
          <input type="number" id="sleep_time" name="sleep_time" min="0" max="24"
required>
         </div>
        <div><label>How many days during the past 30 days was your mental health not
good? </label>
          <input type="number" id="mental_health" name="mental_health" min="0"
max="30" required>
        </div>
        <div><label>How many days during the past 30 days was your physical health not
```

```
good? </label>
            <input type="number" id="physical_health" name="physical_health" min="0"
max="30" required>
          </div>
        </div>
      </div>
    </fieldset>
    <fieldset>
      <legend>
        <h3>Health Information</h3>
      </legend>
      <div  class="personal-details">
        <div>
          <div><label>Did you do exercise during the past 30 days other than their regular
job?</label>
            <select id="physical_activity" name="physical_activity">
            <option value="Yes">Yes</option>
            <option value="No">No</option>
            </select>
          </div>
          <div><label>Blood Pressure(mmHg):</label>
            <input type="number" id="bp" name="bp">
          </select>
        </div>
        <div><label>HMaximum Heart Rate(bpm):</label>
          <input type="number" id="heart_rate" name="heart_rate">
        </div>
        <div><label>Serum Cholesterol(mg/L):</label>
          <input type="number" id="Serum_level" name="Serum_level">
```

```
</div>
<div><label>High Fasting Blood Sugar:</label>
 <input type="number" id="blood_sugar" name="blood_sugar">
</div>
 <div><label>Have you ever had a stroke?</label>
  <select id="stroke" name="stroke">
  <option value="Yes">Yes</option>
  <option value="No">No</option>
  </select>
 </div>
 <div><label>Have you ever had a diabetes?</label>
  <select id="diabetic" name="diabetic">
  <option value="Yes">Yes</option>
  <option value="No">No</option>
  </select>
 </div>
 <div><label>Have you ever had an asthma?</label>
  <select id="asthma" name="asthma">
  <option value="Yes">Yes</option>
  <option value="No">No</option>
  </select>
 </div>
 <div><label>Have you ever had a skin cancer?</label>
  <select id="skin_cancer" name="skin_cancer">
  <option value="Yes">Yes</option>
  <option value="No">No</option>
  </select>
 </div>
 <div><label>Have you ever had a kidney disease?</label>
```

```
            <select id="kidney_disease" name="kidney_disease">
            <option value="Yes">Yes</option>
            <option value="No">No</option>
            </select>
          </div>
        </div>
      </div>
    <button type="submit" href="/">Submit</button>
    </fieldset>
  </form>
  </div>
 </body>
</html>
```

Style.css

```
html {
  min-height: 100%;
  }
  body{
    min-height: 100%;
    background: white;
  }
  body, div, form, input, select, p {
  padding: 0;
  margin: 0;
  outline: none;
  font-family: Roboto, Arial, sans-serif;
  font-size: 14px;
  color: darkgoldenrod;
```

```
}
h1 {
color: 'black;';
margin: 0;
font-weight: 600;
text-align: center;
}
h3 {
margin: 12px 0;
color: darkgoldenrod;
}
.main-block {
justify-content: center;
align-items: center;
}
img {
  padding-right: 0px;
  padding-top: 190px;
  width: 37%;
  float: right;
}
form {
background-color: 'white;';
width: 900px;
height: 900px;
margin: 0;
padding: 30px;
}
fieldset {
```

```
border: none;
border-top: 1px solid #ffd700;
}
.account-details, .personal-details {
display: flex;
flex-wrap: wrap;
justify-content: space-between;
}
.account-details >div, .personal-details >div >div {
display: flex;
align-items: center;
margin-bottom: 10px;
}
.account-details >div{
width: 100%;
}
.personal-details >div >div, input, label {
width: 100%;
}
label {
color: black;
padding: 0 7px;
text-align: right;
vertical-align: middle;
}
input {
padding: 5px;
vertical-align: middle;
}
```

```
    .checkbox {
    margin-bottom: 10px;
    }
    select, .children, .gender, .bdate-block {
    width: calc(100% + 26px);
    padding: 5px 0;
    }
    select {
    background: transparent;
    }
    #smoking, #general_health, #alcohol, #mental_health, #physical_health, #sleep_time{
      width: 20%;
    }
    #physical_activity, #diff_walking, #stroke, #diabetic, #asthma, #skin_cancer,
#kidney_disease{
      width: 20%;
    }
    button {
    width: 100%;
    padding: 10px 0;
    margin: 10px auto;
    border-radius: 5px;
    border: none;
    background: #ffe866;
    font-size: 14px;
    font-weight: 600;
    color: black;
    }
    button:hover {
```

```
background: #ffd700;
}
@media (min-width: 568px) {
.account-details >div{
width: 50%;
}
.personal-details >div {
width: 100%;
}
label {
width: 40%;
}
input {
width: 60%;
}
select, .children, .gender, .bdate-block {
width: calc(60% + 16px);
}
}
```

pie.html

```
<!DOCTYPE html>
<html lang="en-US">
  <head>
    <title>piechart</title>
    <style>
      ::after,*::before{
        margin: 0px;
        padding: 0px;
```

```
            box-sizing: border-box;
        }
    </style>
</head>
<div id="piechart"></div>

<script type="text/javascript"
src="https://www.gstatic.com/charts/loader.js"></script>

<script type="text/javascript">
    // Load google charts
    google.charts.load('current', {'packages':['corechart']});
    google.charts.setOnLoadCallback(drawChart);

    // Draw the chart and set the chart values
    function drawChart() {
        var pieValue1 = ({heart_risk_rate}/100)*14;
        var pieValue2 = pieValue1-14;
        var data = google.visualization.arrayToDataTable([
            ['Task', 'Hours per Day'],
            ['Low Risk', pieValue2],
            ['High Risk', pieValue2],
        ]);

        // Optional; add a title and set the width and height of the chart
        var options = {'title':'Probablity of Getting Heart Diseases', 'width':550,
'height':400};
        options.colors=['green','red']
        // Display the chart inside the <div> element with id="piechart"
```

```
            var chart = new
google.visualization.PieChart(document.getElementById('piechart'));
            chart.draw(data, options);

        }
    </script>

<body>

</body>
</html>
```

Prediction.html

```
<!DOCTYPE html>
<html>
<head>
<style>
body, html {
  height: 100%;
  margin: -14px;
}
.bg {
  background-image: url("static/background.png");
  height: 100%;
  background-position: center;
  background-repeat: no-repeat;
  background-size: cover;
}
h1 {
```

```
  color: #0e0e3f;
  font-size: 50px;
  padding-top: 200px;
  padding-right: 700px;
  text-align: center;
}
h3 {
  color: #d2143a;
  text-align: center;
  padding-right: 700px;
  font-size: 100px;
}
.row{
  display: flex;
}
.col{
  padding:1em;
}
</style>
</head>
<body>
  <div class="row">
    <div class="col">
      <iframe src="pie.html" height="500" width="500" frameborder="0"
scrolling="no"></iframe>
    </div>
    <div class="col">
      <div class="bg">
        <h1>{{name}}, </br> your heart disease risk rate is:</h1>
```

```
      <h3>{{heart_risk_rate}}</h3>
    </div>
   </div>
  </div>
</body>
</html>
```

7.4 Feature 4

main.py

```
from flask import Flask, render_template, request
from prediction_utils import *

app = Flask(__name__)

@app.route('/')
def home():
        return render_template('index.html')

@app.route('/prediction', methods=["GET", "POST"])
def get_prediction():
        name            = request.values.get('name')
        sex             = request.values.get('sex')
        age             = request.values.get('age')
        race            = request.values.get('race')
        height          = request.values.get('height')
        weight          = request.values.get('weight')
        smoking         = request.values.get('smoking')
        alcohol         = request.values.get('alcohol')
        general_health  = request.values.get('general_health')
```

```
        sleep_time       = request.values.get('sleep_time')
        mental_health    = request.values.get('mental_health')
        physical_health  = request.values.get('physical_health')
        physical_activity = request.values.get('physical_activity')
        bp               = request.values.get('bp')
        heart_rate       = request.values.get('heart_rate')
        serum_level      = request.values.get('serum_level')
        blood_sugar      = request.values.get('blood_sugar')
        stroke           = request.values.get('stroke')
        diabetic         = request.values.get('diabetic')
        asthma           = request.values.get('asthma')
        skin_cancer      = request.values.get('skin_cancer')
        kidney_disease   = request.values.get('kidney_disease')

        model_prediction = predict(sex, age, race, height, weight, smoking, alcohol,
general_health, sleep_time, mental_health, physical_health, physical_activity, bp,
heart_rate, serum_level, blood_sugar, stroke, diabetic, asthma, skin_cancer,
kidney_disease)
        return render_template('prediction.html', name=name,
heart_risk_rate=f"{model_prediction}%")

app.run(debug=True)
```

prediction_utils.py

```
import pandas as pd
import pickle

def encode_data(race, diabetic):
        race_encoder    = {'American Indian/Alaskan
```

```
Native':0,'Asian':1,'Black':2,'Hispanic':3,'Other':4,'White':5}
        diabetic_encoder = {'No':0, 'No, borderline diabetes':1, 'Yes':2, 'Yes (during
pregnancy)':3}

        encoded_race     = race_encoder[race]
        encoded_diabetic = diabetic_encoder[diabetic]

        return encoded_race, encoded_diabetic

def min_max_scaling(age, bmi, physical_health, mental_health, sleep_time, bp,
heart_rate):
        MAX_BMI            = 94.85
        MIN_BMI            = 12.02
        MAX_SLEEP_TIME     = 24.0
        MIN_SLEEP_TIME     = 1.0
        MAX_PHYSICAL_HEALTH, MAX_MENTAL_HEALTH = 30.0, 30.0
        MIN_PHYSICAL_HEALTH, MIN_MENTAL_HEALTH = 0.0, 0.0
        MAX_BP                              = 120
        MIN_BP                              = 90

        scaled_bmi                  = (bmi - MIN_BMI) / (MAX_BMI - MIN_BMI)
        scaled_physical_health = (int(physical_health) - MIN_PHYSICAL_HEALTH) /
(MAX_PHYSICAL_HEALTH - MIN_PHYSICAL_HEALTH)
        scaled_mental_health   = (int(mental_health) - MIN_MENTAL_HEALTH) /
(MAX_MENTAL_HEALTH - MIN_MENTAL_HEALTH)
        scaled_sleep_time      = (int(sleep_time) - MIN_SLEEP_TIME) /
(MAX_SLEEP_TIME - MIN_SLEEP_TIME)
        scaled_heart_rate      = (int(heart_rate)-int(age))
        scaled_bp                   = (int(bp)-MIN_BP)/(MAX_BP-MIN_BP)
```

```
    return scaled_bmi, scaled_physical_health, scaled_mental_health,
scaled_sleep_time, scaled_bp, scaled_heart_rate

def load_model():
    file = open("C:\heart\Flask\saved_model_lgbm.h5", "rb")
    model = pickle.load(file)
    return model

def predict(sex, age, race, height, weight, smoking, alcohol, general_health, sleep_time,
mental_health, physical_health, physical_activity, bp, heart_rate, serum_level,
blood_sugar, stroke, diabetic, asthma, skin_cancer, kidney_disease):

    gen_health_dict = {"Poor":0, "Fair":1, "Good":2, "Very good":3, "Excellent":4}
    age_category_dict = {"18-24":0, "25-29":1, "30-34":2, "35-39":3, "40-44":4, "45-49":5,
"50-54":6, "55-59":7, "60-64":8, "65-69":9, "70-74":10, "75-79":11, "80 or older": 12}
    bmi             = int(weight) / ((int(height)/100) **2)

    smoking          = 1 if smoking == "Yes"  else 0
    alcohol          = 1 if alcohol == "Yes"  else 0
    stroke           = 1 if stroke == "Yes"  else 0
    physical_activity = 1 if physical_activity == "Yes"  else 0
    asthma           = 1 if asthma == "Yes"  else 0
    kidney_disease    = 1 if kidney_disease == "Yes"  else 0
    skin_cancer       = 1 if skin_cancer == "Yes"  else 0
    sex              = 1 if sex == "Female"  else 0
    general_health    = gen_health_dict[general_health]
    age              = age_category_dict[age]
    race, diabetic    = encode_data(race, diabetic)
```

```
    bmi, physical_health, mental_health, sleep_time, bp, heart_rate, serum_level,
blood_sugar = min_max_scaling(age, bmi, physical_health, mental_health, sleep_time,
bp, heart_rate, serum_level, blood_sugar)

    df_dict = {"BMI": bmi, "Smoking":smoking, "AlcoholDrinking":alcohol, "Stroke":
stroke, "PhysicalHealth": physical_health, "MentalHealth":mental_health, "Sex":sex,
"AgeCategory":age, "Race":race, "Diabetic":diabetic, "PhysicalActiviy":physical_activity,
"GenHealth":general_health, "SleepTime":sleep_time, "Asthma":asthma,
"KidneyDisease":kidney_disease, "SkinCancer":skin_cancer}
        temp_df = pd.DataFrame(df_dict, index=[0])
        temp_df = temp_df.iloc[0]
        model = load_model()
        predict_proba = model.predict_proba([temp_df])[0][1] * 100

        return round(predict_proba, 2)
```

8.TESTING

8.1 Test Cases

- Verify the username and password entered by the user.
- Verify and validate all the data entered by the user.
- Validate the personal information and detailed information about their health in the home page.
- Verify that the image is visible to the user.

9. ADVANTAGES

- Heart disease is one of the biggest causes of morbidity and mortality among the population of the world.
- Prediction of cardiovascular disease is regarded as one of the most important subjects in the section of clinical data analysis because the amount of data in the healthcare industry is huge.
- This project is to check whether the patient is likely to be diagnosed with any cardiovascular heart diseases based on their medical attributes such as gender, age, chest pain, fasting sugar level, etc.
- Cost effective for patients.
- Reduce the time complexity of doctors.

DISADVANTAGES

- Prediction of cardiovascular disease results is not accurate.
- Data mining techniques does not help to provide effective decision making.
- Cannot handle enormous datasets for patient records

10. CONCLUSION

The leading global cause of mortality and morbidity is heart infection. This study developed a brand-new method of visualising the socio-demographic parameters that affect the prevalence of heart disease. The results suggest that the government and other relevant parties should put more effort into implementing appropriate policy measures to lower the prevalence of heart disease and related risk factors. Additionally, HDV offers an interactive visualisation system that may be used to identify significant trends in the study of heart disease and encourage the implementation of pertinent initiatives and programes to increase public safety. The resource demonstrates the

application's value in identifying trends using relevant health status, building collaboration in a fluid setting, and providing a transparent overview of heart disease progression trajectories parts on consumer and constraint views. To improve the data being investigated in our future study, the suggested visual analytics utility must be used in conjunction with an NLP-based knowledge acquisition technique. This enhanced data can then be fed into a deep-learning system that can accurately anticipate future event patterns. Because viewers can observe the training of HMMs, HDVis will be improved. Additionally, we like to assess the platform's suitability for use with many synchronized views using reliable user research surveys.

11.FUTURE SCOPE

Heart stroke and vascular disease are the major cause of disability and premature death. Chest pain is the key to recognize the heart disease. In this work, the heart diseases are predicted by considering major factors with four types of chest pain. K-means clustering is one of the simplest and popular unsupervised machine learning algorithms. Here the datasets are clustered and based upon the clusters the happening of chest pain is predicted. The role of exploratory data using IBM COGNOS provided a visual appealing and accurate clustering experience.The scope of this project is to check whether the patient is likely to be diagnosed with any cardiovascular heart diseases based on their medical attributes such as gender, age, chest pain, fasting sugar level, etc.

12. APPENDIX

DemoLink:

https://drive.google.com/file/d/13igilESajbWUmy0_Xh80SDqQ1O7CAJky/view?usp=sharing